

10-07-2015 Reconciliation Telecon Minutes

Date

07 Oct 2015

Attendees

- Dana Shum

Discussion Items

- Roundtable of reconciliation status from each DAAC**
- EntryID future:** In order to allow GCMD and CMR clients to work with data in a more consistent manner, we are proposing modifying the creation of the Entry ID in the metadata. UMM-C and DIF 10 will be replacing the Entry ID element in the metadata with elements ShortName and Version. However the element Entry ID will still exist through the CMR API and will be created internally by the CMR based on the ShortName, and if applicable, the version. The ShortName/Version elements will move towards being unique across the CMR, similar to how entry id is currently unique in the GCMD.

Example:

- ShortName = Hello
- Version = World
- CMR Created EntryId = Hello_V:World

DIF 9 Users: The CMR will use the DIF 9 Entry ID for the CMR's Entry ID since DIF 9 does not have a version field.

Only action for the providers is to ensure shortname and version are filled out on DIF and ECHO collections and to understand that the CMR would be overwriting anything in the entryid field. Additionally, the CMR team will be reaching out to select DAACs for those collections that currently do not have a unique shortname.

- Most Recent Version:** A problem that the CMR is trying to solve is being able to identify the most recent version of a collection. We have looked at two approaches (sorting the version string and comparing InsertTime / MetadataCreationTime), but haven't found a solid approach yet. Would like to ask for suggestions.

Ops providers that had collections with more than one version for a given short name

Provider	Number Applicable to Simple Version Sort?	Non-conforming Example (Short name : versions)
ASF	3/3 (100%)	
GSFCS4PA	99/104 (95%)	GLDAS_NOAH025_M : 001, 020, 2.0
LARC	73/73 (100%)	
LPDAAC_ECS	77/94 (82%)	MYD11_L2 : 3, 4, 41, 5, 6
NSIDCV0	1/3 33%	NSIDC-0080 : 1, Undefined
NSIDC_ECS	53/76 (70%)	AQ3_WKSM : 003, 004, 2

From this we can see that we cannot rely on version-id for all providers. For the others it might be possible to use the insert time provided in the metadata (*not* the time of ingest into the CMR). The next table summarizes the relationship between version and insert time for some of the collections for which simple string sorting of version-id does not work:

Insert times for non-conforming collections by version-id

Provider	Short Name	Version	Insert Time
GSFCS4PA	GLDAS_NOAH025_M	001	2007-08-17T00:00:00.000Z
GSFCS4PA	GLDAS_NOAH025_M	020	2013-09-24T00:00:00.000Z
GSFCS4PA	GLDAS_NOAH025_M	2.0	2015-05-28T00:00:00.000Z
LPDAAC_ECS	MYD11_L2	3	2003-02-08T20:51:55.323Z
LPDAAC_ECS	MYD11_L2	4	2004-01-08T11:06:38.813Z
LPDAAC_ECS	MYD11_L2	41	2008-11-12T14:12:13.726Z *
LPDAAC_ECS	MYD11_L2	5	2005-06-14T09:33:38.526Z
LPDAAC_ECS	MYD11_L2	6	2010-10-13T07:39:42.286Z
NSIDCV0	NSIDC-0080	Undefined	2013-10-31T12:00:00.000Z
NSIDCV0	NSIDC-0080	1	2015-09-11T12:00:00.000Z
NSIDC_ECS	AQ3_WKSM	2	2013-11-25T13:13:02.217Z
NSIDC_ECS	AQ3_WKSM	003	2014-08-21T15:09:24.714Z
NSIDC_ECS	AQ3_WKSM	004	2015-07-27T13:49:32.599Z

* MYD11_L2 from provider LPDAAC version 41 (presumably 4.1) has insert time after version 5 but before version 6.

Action Items

